

THE CAMBRIDGE STATISTICS DISCUSSION GROUP

Thursday 20th February 2014 7:15 for 7:45

Tom Ap Rees Lecture Theatre, Department of Plant Sciences,
Downing Site, Cambridge, CB2 3EA

High-dimensional variable selection

Richard Samworth
Statistical Laboratory

Abstract: We live in the era of Big Data, where it is often relatively cheap to collect and store vast quantities of data. While this presents unprecedented opportunities to improve our scientific understanding and quality of life, there are also substantial statistical challenges. One of the most important of these is variable selection, and it is a topic that has received enormous attention among statisticians over the last 15 years or so. I will describe some methods that have been proposed, and outline some of my own work, in collaboration with Rajen Shah, that aims to improve the performance of any existing variable selection algorithm through Complementary Pairs Stability Selection.

Speaker: Richard Samworth obtained his PhD in Statistics from the University of Cambridge in 2004. Following a research fellowship at St John's College, Cambridge, he was appointed to a lectureship in Statistics at the Statistical Laboratory in Cambridge in 2005. He was promoted to a readership in 2010 and to a full professorship from October 2013. He remains a fellow of St John's College, and currently holds an EPSRC Early Career Fellowship (worth GBP 1.2M) for five years from December 2012. His main research interests are in nonparametric and high-dimensional statistics. Particular topics include shape-constrained density and other nonparametric function estimation problems, nonparametric classification, clustering and regression, Independent Component Analysis, the bootstrap and high-dimensional variable selection problems. He was awarded the Royal Statistical Society Research prize (2008), a Leverhulme Research Fellowship (2011) and the Royal Statistical Society Guy Medal in Bronze (2012). He currently serves as an Associate Editor for the *Annals of Statistics*, the *Journal of the Royal Statistical Society Series B*, *Biometrika* and *Statistica Sinica*.

Directions: The Department of Plant Sciences is located on the Downing site situated between Downing Street and Tennis Court Road (Please see the map on the reverse). Go through the second door you come to in the building immediately on your left as you enter from Tennis Court Road. Once inside the building follow the arrows. Refreshments will be served in the first floor kitchen. It is not possible to park on the Downing site in the evening. The nearest car park is at Lion Yard. Arrivals after 7-45pm can gain admittance by contacting the secretary on 07761769436.

Provisional Next Meetings (Please note switch in March and April talks – apologies for any inconvenience):

27th March – Nick Galwey (GlaxoSmithKline) on 'Working with Epidemiologists'.

30th April – Peter Fretwell (British Antarctic Survey).

14th October – Adrian Mander (MRC Biostatistics).

17th November – Trevor Lewis (TLwise Consulting Limited).

Supper: Some members eat regularly in the University Centre before each meeting at **6pm**. Feel free to join them.

Subscriptions: of 1 pound are now due for attending the 2013-2014 session.

Secretary: Peter Watson, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF;
telephone 01223 355294 Extension 801; E-mail peter.watson@mrc-cbu.cam.ac.uk

Slides and .mp3 files of old talks: <http://www.mrc-cbu.cam.ac.uk/people/peter.watson/csdg.html>

PLANT SCIENCES

